Nucleic Acids Research, 2006, Vol. 34, No. 11 ***e76***
doi:10.1093/nar/gkl391

Primer Extension Enrichment Reaction (PEER): a new subtraction method for identification of genetic differences between biological specimens

Lilia Ganova-Raeva*, Xinjian Zhang, Fengli Cao, Howard Fields and Yury Khudyakov

Centers for Disease Control and Prevention, National Center for Infectious Diseases,
Division of Viral Hepatitis/Laboratory Branch, Atlanta, GA 30329, USA

Received February 1, 2006; Revised April 20, 2006; Accepted May 8, 2006

ABSTRACT

We developed a conceptually new subtraction strategy for the detection and isolation of target DNA and/or RNA from complex nucleic acid mixtures, called Primer Extension Enrichment Reaction (PEER). PEER uses adapters and class IIS restriction enzymes to generate tagged oligonucleotides from dsDNA fragments derived from specimens containing an unknown target ('tester'). Subtraction is achieved by selectively disabling these oligonucleotides by extension reaction using ddNTPs and a double stranded DNA template generated from a pool of normal specimens ('driver'). Primers that do not acquire ddNTP are used to capture and amplify the unique target DNA from the original tester dsDNA. We successfully applied PEER to specimens containing known infectious agents (Hepatitis B Virus and Walrus Calicivirus) and demonstrated that it has higher efficiency than the best comparable technique. The strategy used for PEER is versatile and can be adapted for the identification of known and unknown pathogens and mutations, differential expression studies and other applications that allow the use of subtractive strategies.

INTRODUCTION

The isolation of unique fragments from complex nucleic acid mixtures without prior knowledge of their sequence or abundance is technically challenging. The earliest approach to the identification of target nucleic acids of unknown sequence, which involved direct screening of large libraries, is inefficient and labor intensive (1). In the last decade more robust techniques have been developed. They can be divided into two categories: (i) subtractive approaches such as differential display (2,3), representational differences analysis (RDA) (4) and its variations (5,6), differential subtraction chain (DSC) (7), Selective Amplification via Biotin- and Restriction-mediated Enrichment (SABRE) (8), Suppression Subtractive Hybridization (SSH) (9) and (ii) high-throughput methods like sequencing-by-hybridization (10), microarrays (11,12) and massive parallel sequencing (13,14). Integration of subtractive approaches into high-throughput methods (15) can also be used for the identification of unknown sequences.

All subtractive approaches are based on molecular comparison of two specimens: 'tester', a specimen that is suspected to contain the unknown target of interest, and 'driver', a specimen that is a perfect genetic match for the tester but is believed not to contain the target. Subtractive methods are often used in molecular studies because of their relative simplicity and high efficiency. Among the subtractive techniques, RDA and the closely related SSH are the most popular and have been successfully used to recover unknown sequences. SSH can enrich a target gene $\sim 3 \times 10^3$ (1,9) and was used to find a new calicivirus in walrus (16); GBV-A and -B viruses were found by RDA (17,18). Although impressive, this performance is not sufficient for the detection of an infectious agent that may be present at only a few copies in the specimen of interest. Some limitations of the subtractive approaches are the requirements for perfect hybridization and an abundance of ideally matched driver, which makes them intrinsically biased against single-stranded, low-copy-number molecular species (17,19). Enrichment of the target of interest is achieved by hybridization between long and sometimes heterogeneous population of DNA fragments (e.g. cDNA generated by random priming). If present in low numbers, such molecules have little chance to form complete hybridization products after denaturing. The fragment length can also compromise the specificity by creating background cross-hybridization. In addition, many subtraction approaches rely on the presence of a poly-A tail to generate the starting material and consequently are not suitable when working with DNA or RNA that is not polyadenylated.

The high-throughput approaches are limited by cost, some are not suitable for use with small sample volumes and others require prior knowledge of the target sequence.

---

*To whom correspondence should be addressed. Tel: +1 404 639 1158; Fax: +1 404 639 1563; Email: lkg7@cdc.gov

© 2006 The Author(s).
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commerical use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Primer Extension Enrichment Reaction (PEER) method developed in the current study is a new subtraction techniques that is built on two novel principles: the use of tester DNA to generate both PCR primer and template, and the selective inactivation of primers containing sequences common to the tester and driver to ensure preferential amplification of templates that contain sequences unique to the tester.

MATERIALS AND METHODS

PEER outline

A general outline of PEER is presented in Figure 1. Total nucleic acid (NA) is extracted from a tester and a driver specimen and used in a modified SMART cDNA protocol (Clontech, Palo Alto, CA) to generate double stranded DNA (dsDNA) with two different sets of primers for the tester and one set for the driver. The product is referred to as dsDNA to distinguish it from cDNA since it was generated from total nucleic acid instead from RNA alone (Figure 1, A1–A4). Tester 1 dsDNA material is converted into small fragments by extensive endonuclease cleavage and then tagged by ligation to a specially designed adapter. The 3′ end of the adapter incorporates a recognition site for a class IIS restriction endonuclease (20). After the ligation the fragments are cleaved with the IIS enzyme to create oligos with unique sequence at the 3′ end derived from the tester and a 5′ end derived from the adapter (Figure 1, B1–B3). These adapter-tagged oligos are annealed to driver dsDNA template and extended in the presence of biotinylated ddNTPs. All oligos that prime a reaction from the driver template can acquire biotinylated ddNTP. This event blocks any further extension and allows the removal of the biotinylated molecules from the reaction by use of streptavidin-coated magnetic beads. Primers that share driver sequences are blocked and removed leaving only primers with unique sequences that can only be found in the tester (Figure 1, C1–C3). In the presence of Tester23 dsDNA and dNTPs, these oligonucleotides can prime an extension reaction from the fragments unique to the tester (target capture). This step converts the tagged primers into DNA templates suitable for PCR amplification by oligonucleotides containing only the adapter sequences or in combination with T2PCR or T3PCR oligos. The last step in PEER is a standard PCR amplification with primers containing only adapter and T2PCR/T3PCR sequences that can be used without any molarity restrictions. The final step is expected to generate collection of fragments of different sizes (Figure 1, D1–D3).

Blocking experiments

To test and optimize the blocking efficiency we tested a variety of polymerases: Vent (exo-) polymerase and Deep Vent (exo-) polymerase (New England Biolabs, Inc., Beverly, MA); Tth polymerase and Tfl polymerase (Promega, Madison, WI), Thermo Sequenase™ (Amersham Pharmacia Biotech, Inc., Piscataway, NJ) and *Taq* Polymerase (Roche, Indianapolis, IN); and a range of nucleotide concentrations using pB6 as template and 50 pmol each SK and T7 generic primers as shown in Figure 2. After 45–55 cycles of extension in the presence of the ddNTPs-bio, one-fifth of the product was transferred to a fresh PCR tube, supplemented with *Taq* DNA polymerase, buffer and dNTPs to a final reaction volume of 50 μl, and subjected to 35 cycles at conditions suitable for the amplification of the particular insert in this clone.

Target-capture experiments

We used oligonucleotides AT7 and ASK (Table 1) as capture primers. They were mixed in a master stock with concentrations 8–16 amol of template (pB6) and 100–500 fmol of primer (AT7/ASK) to approximate the actual primer: template ratio expected to be generated from a target of similar size by the PEER protocol. Serial 2-fold dilutions of the stocks were subjected to 50 PCR cycles (95°C for 20 s, 45°C for 30 s, 72°C for 60 s). One-tenth of each product was subjected to a second round of PCR with 100 pmol of adapter primer. As a positive PCR control, the same template was amplified with the generic SK and T7 primers.

dsDNA and PEER product enrichment analysis

All dsDNA products, as generated in step A4 (Figure 1), were tested for the presence of the desired target by PCR with virus specific primers to confirm their initial titers before performing any enrichment procedures, e.g. dsDNA generated from the serum containing 3.6×10^8 IU virus was serially diluted 10-fold and each dilution was used as a PCR template with HBV specific primers. The sensitivity of our modified (21) HBV specific PCR is <10 copies/ml. The same procedure was used for all different serum dilutions.

Spot hybridizations were performed with all dsDNA materials and PEER products to confirm the presence of the target of interest and to assess the level of enrichment in the following manner: 10% of all dsDNA and PEER products were subjected to serial 2-fold dilutions, denatured, spotted on positively charged nylon membranes, ultraviolet cross-linked and probed by Southern hybridization (Roche DIG-labeling and hybridization kit, sensitivity—100 fg) with digoxigenin-labeled HBV genome or WCV genome, respectively. The hybridization was done as described by the manufacturer and the last stringency wash was at 45°C. The positive hybridization control was genomic DNA from the corresponding target virus with known concentration also spotted in 2-fold serial dilutions. The initial concentration of the tested products was calculated by measuring A_{260}. The enrichment was calculated by dividing the amount of targets found after PEER by the amount found before the enrichment calculated as percent of the applied DNA.

In addition to the spot hybridization, the dsDNAs and PEER products were cloned in separate libraries using pTAdvantage vector (Clontech, Palo Alto, CA) and *Escherichia coli* Top 10F′ electrocompetent cells (Clontech). Up to 2000 clones were randomly isolated and sequenced from the high titer libraries (3×10^8 and 3×10^7). Up to 400 clones were isolated randomly from each PEER library and sequenced. The clones selected for sequencing were subject to PCR with generic vector primers and the resulting fragments purified on BioRobot8000 using the QIAquick 96 PCR Biorobot kit (Qiagen, Inc., Valencia, CA). Sequencing was done on ABI3100 DNA Sequencer (Applied Biosystems, Foster City, CA) with Bid Dye v3.1 chemistry. All sequences

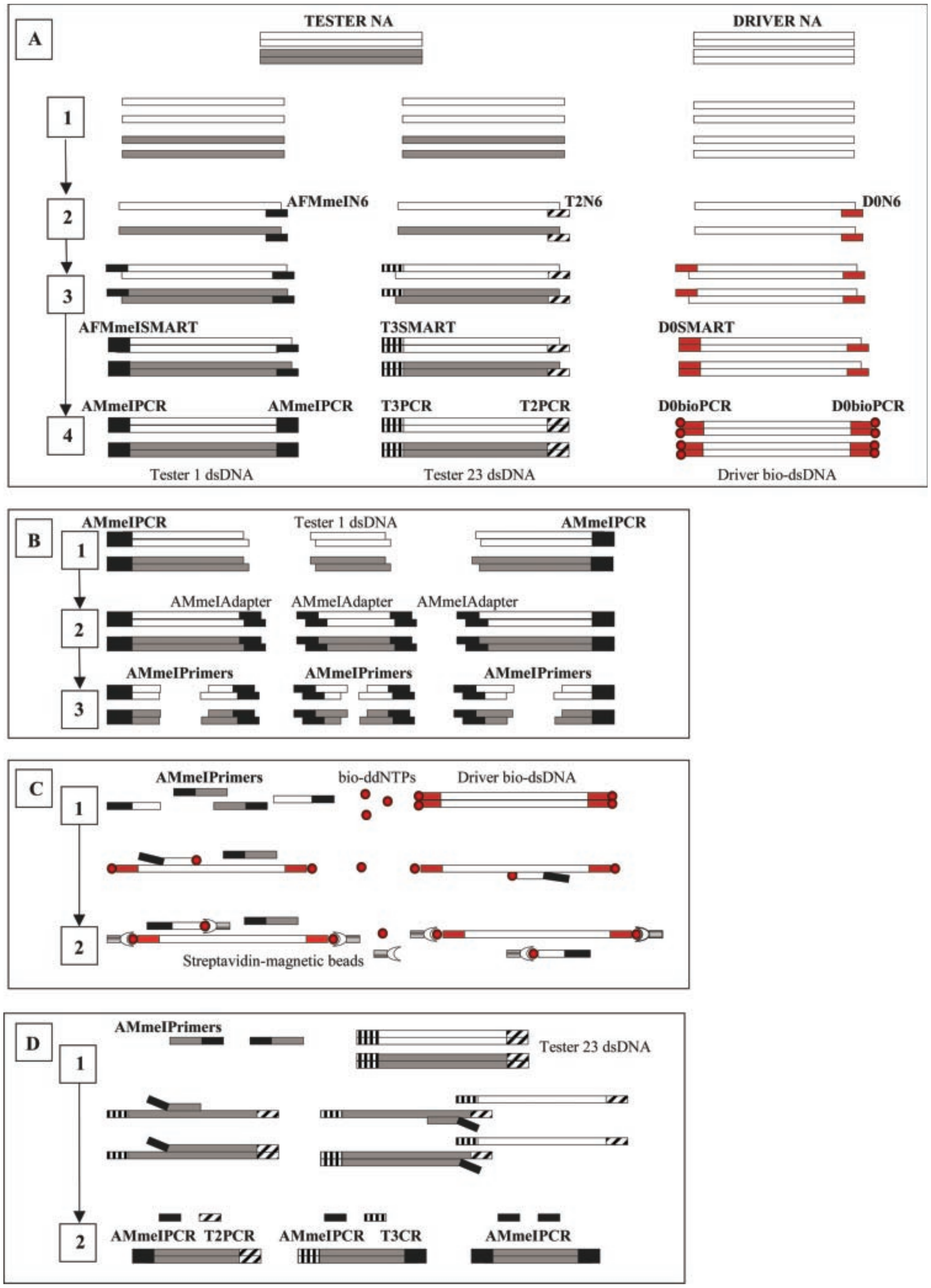
A
TESTER NA
DRIVER NA
1
2
AFMmeIN6
T2N6
D0N6
3
AFMmeISMART
T3SMART
D0SMART
4
AMmeIPCR
AMmeIPCR
T3PCR
T2PCR
D0bioPCR
D0bioPCR
Tester 1 dsDNA
Tester 23 dsDNA
Driver bio-dsDNA
B
AMmeIPCR
Tester 1 dsDNA
AMmeIPCR
AMmeIAdapter
AMmeIAdapter
AMmeIAdapter
AMmeIPrimers
AMmeIPrimers
AMmeIPrimers
C
AMmeIPrimers
bio-ddNTPs
Driver bio-dsDNA
Streptavidin-magnetic beads
D
AMmeIPrimers
Tester 23 dsDNA
AMmeIPCR
T2PCR
AMmeIPCR
T3CR
AMmeIPCR

Figure 1. Primer Extension Enrichment Reaction (PEER). (**A**) Generation of dsDNA from total Nucleic Acid. (1) Tester NA (white and gray rectangle) is split in two aliquots and denatured; Driver NA (white rectangle) is denatured as well. (2) Single strands are reverse transcribed (RT) by Super Script RT with three different primers—AFMmeIN6* for the first Tester aliquot, T2N6 (diagonal fill rectangle) for the second aliquot and D0N6 (red rectangle) for the Driver. (3) The reverse transcriptase switches templates and copies the annealed SMART primers (SMART technology, Clontech). (4) The RT products are amplified with Advantage2 Polymerase to yield Tester1 dsDNA with primers AMmeIPCR (black rectangle), Tester23 dsDNA with T3PCR (vertical fill rectangle) and T2PCR (diagonal fill rectangle) and Driver bio-dsDNA with D0bioPCR biotinylated at the 5′ end (red rectangle with red circle). (**B**) Processing of Tester1 dsDNA. (1) The DNA is cleaved by a cocktail of restriction enzymes that leave 3′ GC protruding ends. (2) The ends are treated with the Klenow fragment of DNA Polymerase I in the presence of dCTP only and then ligated to AMmeIAdapter. (3) The tagged fragments are cut to uniform size by MmeI to create multiple AMmeIPrimers. (**C**) Blocking reaction. (1) AMmeIPrimers generated from Tester1 dsDNA are extended on Driver bio-dsDNA template in the presence of biotinylated ddNTPs (red circles) and ThermoSequenase™. (2) Biotinylated molecules are captured with streptavidin-coated magnetic beads (white crescent with gray bar) and removed from the reaction. (**D**) Retrieval of targets of interest from the Tester23 dsDNA. (1) Capture PCR—AMmeIPrimers that were not blocked and removed in the preceding steps are added to Tester23 dsDNA and in the presence of regular dNTP are annealed and extended to capture the targets of interest. (2) Regular PCR amplification of the capture products with different primer combinations. Black rectangles, primers AFMmeIN6, AFMmeISMART, AMmeIPCR, AMmeIAdapter.

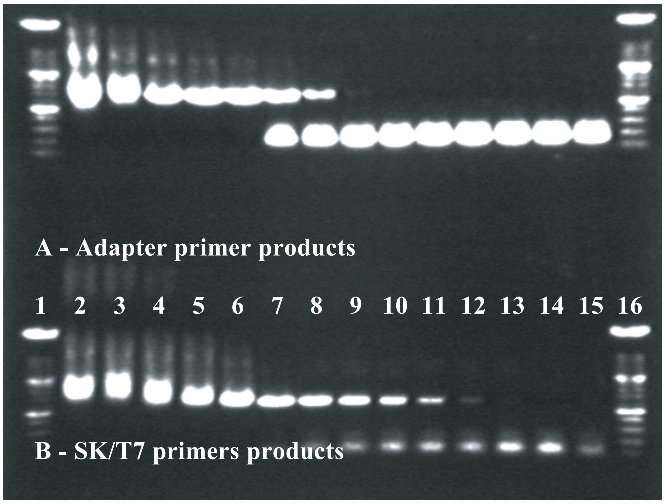

Figure 2. Second-round products of the PEER target-capture experiments. Oligonucleotides AT7 and ASK were used as capture primers on pB6 template mixed as a master stock at a ratio of 8 amol of template to 500 fmol of primer. Totally 12 serial dilutions of this stock in a 1 μg/μl solution of salmon sperm DNA were subjected to 50 PCR cycles (95/20, 45/30 and 72/60 s). For both panels, wells 1 and 16 contain molecular-size markers and wells 14 and 15 are negative controls. (**A**) Of each product, 10 μl was amplified in a second-round PCR with 100 pmol of adapter primer only. Adapter primer reactions yielded amplification products from as little as 0.063 amol of template with as little as 4 fmol of capture primers. (**B**) As a positive PCR control, 10 μl aliquots of the same templates were amplified with the generic SK and T7 primers. These reactions generated the expected product throughout the entire range of dilutions. Wells 2–15 are the second-round products generated on templates from the different template primer concentrations; starting at 8 amol template and 500 fmol capture primers in well 2; 4 amol template/250 fmol primers in well 3 and so on.

were examined by BLAST at the NCBI site (NCBI, http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). Enrichment for the target of interest was calculated by dividing the number of target clones found and confirmed by sequencing in the PEER libraries by the number of targets found and confirmed by sequencing in the 3×10^7 tester library prior to the enrichment.

Test Samples

To represent an RNA virus we used VMK cells infected with a newly discovered calicivirus (16), walrus calicivirus (WCV, positive-sense, single-stranded RNA virus with no DNA stage). VMK cells infected with WCV at 10^2 p.f.u. were used as tester and non-infected VMK cell culture as driver. To represent a DNA virus as tester, hepatitis B virus (HBV), isolate HLD1, derived from an experimentally infected chimpanzee was used at several dilutions of a 3.36×10^8 IU source serum quantified by *RealArt™ HBV LC* (diagnostic limit 3.5 IU, ARTUS-Biotech) real-time PCR on LightCycler1.5 (Roche, Indianapolis, IN). A pool of 16 normal blood donor sera was used as a driver and as a tester diluent. The plasmid pB6, which contains a 650 bp fragment of WCV cloned in pTAdvantage (Clontech), was used as a control template for some preliminary test experiments. The primers designed for use in this study are listed in Table 1.

PEER protocol

Extraction: Total nucleic acid (NA) is extracted from 100 to 200 μl of serum or cell culture using Masterpure Complete kit (Epicenter Biotechnologies, Madison, WI) or High Pure viral nucleic acid extraction kit (Roche) and resuspended in 10 μl of 10 mM Tris (pH 8–8.5).

Modified SMART protocol: 5 μl Of the extracted NA is reverse transcribed (RT) with SuperScript II (Invitrogen, Carlsbad, CA) (22). Two RT reactions are performed for the tester, one using 10 pmol each primer AFMmeIN6 and AFMmeISMART and the other using 10 pmol primers T2N6 and T3SMART. Primers D0SMART and D0N6 are used for the driver reaction. Reaction volumes and conditions are described in the SMART cDNA synthesis protocol (Clontech). After synthesis, the enzyme is heat-inactivated and the product diluted with 40 μl of TE.

PCR amplification: 10 μl Of the RT product is amplified with Advantage 2 Polymerase (Clontech) as recommended in the Smart cDNA protocol and using the corresponding PCR primers (AMmeIPCR for Tester 1, T3PCR and T2PCR for Tester 23 and D0bioPCR for the driver) in triplicate reactions under the conditions suggested by the manufacturer. The amplification parameters are 95°C for 1 min; (95°C for 3 s, 68°C for 3 min) × 28 cycles. The dsDNA is purified on a Qiagen PCR purification column (Qiagen, Inc., Valencia, CA) and eluted in 75 μl of 10 mM Tris (pH 8)

Digestion with restriction endonucleases: 70 μl Of the Tester 1 dsDNA are digested overnight with HpaII, HinP1I, AciI (Roche Molecular Biochemicals, Germany), MaeII (MBI Fermentas Amherst, NY) and TaqI (NEB Ipswich, MA) using 1 μl of each enzyme and TaqI buffer (NEB) at 37°C. After digestion, the enzymes are heat-inactivated; the fragments purified through a QIAquick PCR purification kit and eluted in 55 μl of 10 mM Tris (pH 8).

Klenow treatment: The ends of the fragments are filled in with Klenow polymerase (Roche) in the presence of dCTP for 1 h at 37°C. The enzyme is then heat-inactivated; the reaction mixture purified with QIAquick nucleotide removal column (Qiagen), and the product eluted in 50 μl of 10 mM Tris (pH 8).

Adapter ligation: Double-stranded adapters are prepared by mixing the forward (AFMmeI) and reverse (ARMmeIP) adapter primers (Table 1) at equimolar ratio (200 pmol each), heating to 96°C for 5 min and slowly cooling to room temperature. The adapter (200 pmol) is ligated overnight to 45 μl dCTP-filled-in Tester 1 fragments. The ligation products are purified to remove the T4 ligase and buffer with QIAquick nucleotide removal column and eluted in 55 μl of 10 mM Tris (pH 8).

MmeI digestion: The ligation products are digested with 5 U MmeI (NEB) for 2 h. The cleaved DNA is resolved in 10% polyacrylamide gel, the resulting 50 bp fragment is cut out, isolated from the gel with QIAquick gel extraction kit (Qiagen) and resuspended in 50 μl of 10 mM Tris (pH 8).

Blocking of MmeI-tagged primers: 25 μl Of the fragment is used as primer with 10 μl Driver bio-dsDNA template in the presence of 2.5 mM each ddNTPs-bio (Biotin-11-ddNTPs, NEN® Life Science Products Inc., Boston, MA), 0.025 mM each dNTPs (Roche) and Thermo Sequenase™ (Amersham Pharmacia Biotech, Inc., Piscataway, NJ). The blocking reaction is carried out as follows: 96°C for 3 min; (95°C for 2 s; 55°C for 20 s; 68°C for 20 s) × 55 cycles. The product is purified with QIAquick nucleotide removal kit to remove the excess ddNTPs and eluted in 100 μl of 10 mM Tris (pH 8).

Table 1. Primers used in the PEER study

Primer sequence 5′–3′	Name	Function
MmeI experiments		
AATGCAGACACAGAAGGTCCATCCGAC	AFMmeI	TESTER MmeI adapter forward
P-GGTCGGATGGACCTTCTGTGTCTGC	ARMmeIP[a]	TESTER Mmei adapter reverse
GCTGCAGACACAGAAGGTCCATCCGACNNNNNN	AFMmeIN6	TESTER 1 cDNA
GCTGCAGACACAGAAGGTCCATCCGACGGG	AFMmeISMART	TESTER 1 cDNA
CAGACACAGAAGGTCCATCCGAC	AMmeIPCR	TESTER 1 cDNA PCR
ACACTAGAGCATGCGTCAAGAGAANNNNNN	T2N6	TESTER 23 cDNA
ACACTCCAGGAGGTCAGAAACAACGGG	T3SMART	TESTER 23 cDNA
ACACTAGAGCATGCGTCAAGAGAA	T2PCR	TESTER 23 cDNA PCR
ACACTCCAGGAGGTCAGAAACAAC	T3PCR	TESTER 23 cDNA PCR
AAGCAGTGGTATCAACGCAGAGTANNNNNN	D0N6	DRIVER cDNA
AAGCAGTGGTATCAACGCAGAGTACGCGGG	D0SMART[a]	DRIVER cDNA
Bio-AAGCAGTGGTATCAACGCAGAGTA	D0bioPCR	DRIVER cDNA PCR
BpmI experiments (proof-of-concept)		
ACACTCGAGGAGGTCTGGAGIIIIII	PEER1BpmN6	TESTER 1 cDNA
ACACTCGAGGAGGTCTGGAGGG	PEER1BpmG	TESTER 1 cDNA
AACACTCGAGGAGGTCTGGAG	PEER1BpmAF	TESTER BpmI Adapter Forward
CTCCAGACCTCCTCGAGTGTG	PEER1BpmAR	TESTER BpmI Adapter Reverse
GAGCTGTGGTGAGTTGGTTGGAAIIIIIII	PEERT7N7	TESTER 78 cDNA
AAGCAGAGGCAGCATTGGAGGG	PEERT8G	TESTER 78 cDNA
AGCTGTGGTGAGTTGGTTGG	PEERT7	TESTER 78 cDNA RCR
AGCAGAGGCAGCATTGGAGG	PEERT8	TESTER 78 cDNA RCR
AAGCAGTGGTATCAACGCAGAGTAIIIIII	D0N6	DRIVER cDNA
AAGCAGTGGTATCAACGCAGAGTACGCGGG	D0SMART[b]	DRIVER cDNA
AAGCAGTGGTATCAACGCAGAGTA	D0PCRbio	DRIVER cDNA PCR
Control primers		
AATGCAGACACAGAAGGTCCATCCGAC*TAATACGACTCACTATAGGG*	A*T7*[c]	PEER control primer
AATGCAGACACAGAAGGTCCATCCGAC*GAAACAGCTATGACCATGAT*	A*SK*[3]	PEER control primer

I, 5-nitro indol; N, random base.

[a] P indicates that the oligo was phosphorylated to improve ligation.

[b] According to the SMART cDNA technology (Clontech, Palo Alto, CA).

[c] These primers are not part of PEER but were used to monitor the success of the protocols' steps using a 'control' template as described in Results.

Removal of biotinylated products: The cleaned product is heated to 95°C, and 50 µl of streptavidin-coated magnetic beads were added (SPHERO™ Streptavidin Magnetic Particles from Spherotech, Inc., Libertyville, IL). After 10 min incubation at >60°C, the beads are captured on a magnet rack (Qiagen) and the supernatant removed to a fresh tube taking care that in the process the temperature remains >55°C.

Capture reaction: 50 µl of the supernatant (purified non-blocked primers) are used in a 100 µl capture reaction with 5 µl of the Tester 23 cDNA as template under the following conditions: 95°C for 2 min; (95°C for 20 s; 45°C for 30 s; 72°C for 2 min) × 10 cycles; (95°C for 20 s; 52°C for 30 s; 72°C for 2 min) × 30 cycles; 72°C for 7 min.

PCR: 5 µl Of the capture product is amplified in a 100 µl final reaction volume with primers AMmeIPCR and T2PCR, AMmeIPCR and T3PCR, or AMmeIPCR alone under the following conditions: 95°C for 2 min; (94°C for 10 s; 60°C for 20 s; 72°C for 90 sec) × 30 cycles. The product is quantified, cloned and sequenced.

RESULTS

Target-capture experiments

PEER was designed to find unknown targets at unknown and potentially very low concentrations. To challenge this goal, we conducted experiments aimed at identifing the amount of target DNA in a mixture that can be found and captured using low concentrations of oligonucleotides designed so that the 3′-terminal half matches the template and the 5′-terminal half cannot be found in the template. We also conducted experiments to determine whether this template could be amplified by only the mismatched portion of the capture oligonucleotides as described in Materials and Methods. The controls generated the expected product throughout the entire range of dilutions (Figure 2), and the adapter primer reactions (i.e. PCR with primers whose sequences did not exist in the original template) yielded amplification products from as little as 0.063 amol of template (136 copies) and with as little as 4 fmol of capture primers.

Blocking experiments

To ensure that a large number of primers could be successfully and specifically blocked by di-deoxytermination, we tested a variety of polymerases and a range of nucleotide concentrations using the pB6 template and 50 pmol each of SK and T7 generic primers as described above. The best results, as measured by the absence of product in the reactions to which ddNTPs were added, were achieved with Thermo Sequenase™ (Figure 3). We also observed blocking by Vent (exo-) polymerase (Promega) and *Taq* polymerase (Roche) when the ddNTP: dNTP ratio was 10:1, but Thermo Sequenase™ remained the enzyme of choice because it gave consistent results under all experimental conditions.

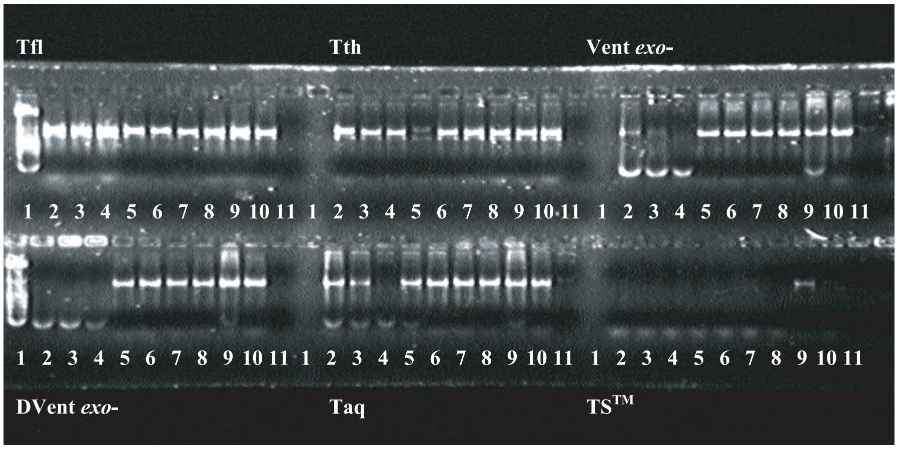

Figure 3. Primer blocking by di-deoxytermination. Well 1 contains 100 bp molecular-size standard (Roche) for Tfl and DVent (exo-) test sets and is empty for the rest of the enzyme sets. Wells 2–11 are the same for each enzyme tested and contain PCR product generated after supplementing 20% of the blocking reaction with fresh *Taq* and 10 mM dNTPs. Well 2, blocked with 10 mM ddNTPs and 10 mM dNTPs; Well 3, blocked with 10 mM ddNTPs and 5 mM dNTPs; Well 4, blocked with 10 mM ddNTPs and 1 mM dNTPs; Well 5, blocked with 10 mM ddNTPs and 0.1mM dNTPs; Well 6, blocked with10 mM ddNTPs and 5 mM dATP and dCTP; Well 7, blocked with 10 mM ddNTPs and 5 mM dGTP and dTTP; Well 8, blocked with10 mM ddNTPs; Well 9, no ddNTPs in the blocking step; Well 10, no dNTPs in the blocking step Well 11, no primers or template.

Table 2. Enrichment for targets of interest by PEER with MmeI Adapters

Tested material		Target presence confirmed by	Hybridization[b]			Sequencing[c]		
Virus	Titer IU	PCR titer[a]	Before (T23)	After (PEER)	Enrichment[d]	Before (T23)	After (PEER)	Enrichment[d]
HBV HLD1	3.36×10^8	10^8	+	NT	NA	3.12% (96)	NT	4.80×10^2
	3.36×10^7	10^7	+	NT	NA	2×10^{-1}% (2000)	NT	na
	3.36×10^6	10^6	NT	NT	NA	2×10^{-2}[e]%	10.60% (122)[f]	5.30×10^2
	3.36×10^5	10^5	6.26×10^{-3}%	3.21%	5.12×10^2	2×10^{-3}[e]%	1.00% (200)	5.00×10^2
	3.36×10^3	10^3	4.76×10^{-5}%	0.781%	1.64×10^4	2×10^{-4}[e]%	0.26% (380)	1.30×10^4
WCV 7240	1.00×10^4	10^3	9.60×10^{-6}%	0.314%	3.27×10^4	2×10^{-4}[e]%	0.89% (112)	4.45×10^4

NT = not tested; The tested material column identifies the virus isolate and the used serum titer.

[a] Aliquots from the dsDNA materials before cloning or enrichment were subjected to 10-fold serial dilutions and the presence of the desired target was confirmed by PCR in all dilutions as indicated (e.g. $\times 10^5$ indicates that the target was amplified from five consecutive dilutions).

[b] Aliquots from the dsDNA materials and PEER products were subjected to serial dilutions, spotted on positively charged nylon membranes and probed with digoxigenin-labeled fragments from the target virus by Southern hybridization. The detected amount of target DNA is shown as percent of the initial aliquot.

[c] Aliquots of the DNA materials were cloned in *E.coli* and the clones were isolated and sequenced at random (the number of sequenced clones is in brackets).

[d] Enrichment was calculated by dividing the values in column 'After (PEER)' by the corresponding values in column 'Before (T23)'; NA = not applicable.

[e] The number is extrapolated from the last available value confirmed by sequencing (i.e. 0.2).

[f] This experiment was done with BpmI adapter.

Enrichment efficiency

We conducted initial proof-of-concept experiments using adapters designed to be compatible with the IIS enzyme BpmI. As a tester, we used serum (HLD1) obtained from an experimentally infected chimpanzee with an initial hepatitis B virus (HBV) titer of 3.36×10^8 IU and as driver and diluent—pooled human sera from normal blood donors. After probing a Tester1 dsDNA library with dig-labeled HBV genome and sequencing we observed that 3.12% of the clones were from HBV (Table 2). This is consistent with previously published observations of target clone frequencies observed in cDNA libraries generated by random amplification (9,18). The cloned PEER products obtained from 100-fold dilution of the tester serum yielded 10.6% HBV clones, which represents an enrichment of 3.39×10^2. Although this value is high, it did not exceed the efficiency of previously published enrichment methods (1).

The final version of the enrichment protocol included a redesign of the cDNA primers and the adapter to accommodate a recognition site for MmeI, a novel IIS restriction enzyme (23) that cleaves 20 bp downstream of its recognition site (24). Digestion with MmeI creates primers with longer unique sequences at their 3′ ends and thus they are expected to have significantly improved primer specificity. To test this modification, we used the same HBV infected serum specimen as the model for a DNA virus and VMK cells infected with WCV at p.f.u. 10^2–10^3 as the model for an RNA virus. We tested all generated dsDNA material for the presence of the desired target by dilution PCR with virus-specific primers. We were able to confirm the presence of viral targets in all expected dilutions (Table 2). These results indicate that we consistently achieved accurate representation of the viral target in all test scenarios. Compared with the corresponding starting dsDNA materials, the PEER products were estimated by hybridization to contain between 500- and 32 700-fold more target molecules (Table 2). To obtain another more accurate measure of the degree of enrichment, we cloned the dsDNA materials and corresponding PEER products to generate paired before and after enrichment libraries. The clones were screened again by hybridization and/or sequencing at random. The sequencing results demonstrated that the enrichment was 5×10^2 when the initial titer of the HBV serum was 3.36×10^5 IU, and 1.3×10^4 when the initial titer of the HBV serum was 3.65×10^3 IU. These numbers agree well with the enrichment calculated from the data obtained by spot hybridization (Table 2). The HBV fragments found in the cDNA library before enrichment came from approximately the 2800–3100 nt position of the genome, whereas the enriched product contained fragments that mapped to positions 100–350 nt, 1400–1500 nt and 2000–2150 nt—all in regions with high occurrences of recognition sites of the restriction enzymes (AciI, HpaII, Hinp1I, TaqI, MaeII) used to generate the primers in our enrichment protocol (Figure 1, B1).

Efficiency of PEER in comparison with SSH

We compared the performance of PEER with that of SSH using WCV isolate 7240 inoculated into VMK cell culture as a tester and VMK cells alone as a perfect driver. We were previously successful in isolating and describing WCV from the same source by SSH. The cDNA materials and PEER products were treated and screened as described above for HLD1. We found that 0.89% of the WCV PEER library clones contained the fragments of interest which translates into enrichment of 4.45×10^4, i.e. more than eight times greater than the enrichment of 5.31×10^3 observed by SSH.

DISCUSSION

PEER exploits unique target sequences by creating primers from the double-stranded material of interest and then using an intact aliquot of the material as a template for amplification. Unlike the SMART protocol we do not use poly-A primers and use total nucleic acid in place of RNA as starting material. The first strand of cDNA is instead created with RT primer that has a random hexamer at the 3′ end and the adapter sequence at the 5′ end. The use of SuperScript™II Reverse Transcriptase, a derivative of M-MLV with DNA Polymerase activity (22), ensures that single-stranded RNA, DNA or RNA:DNA hybrid will be copied into cDNA and enter the enrichment process. Once priming sites are generated on both ends of the fragments the product can be exponentially amplified by SMART PCR to generate dsDNA. This approach maintains the correct representation of all nucleic acids entering the protocol (25) and supplies a renewable source of the target material. Our initial primer design (Table 1) included 5-nitro-indol instead of random bases at the 3′ end of the RT primers. In the later primer design the random bases were favored because we observed a lower efficiency of the PCR step and additional experiments (data not shown) convinced us that the 5-nitro-indol's higher affinity to itself hinders the reaction performance. The PEER protocol can be modified for use with other pre-dsDNA/cDNA procedures. DNase/RNase treatment, filtration, ultracentrifugation, gradient separation, etc. could be recommended depending on the application.

Once generated, the double-stranded material is converted into unique oligonucleotides by extensive endonuclease cleavage (Figure 1, B1) to ensure that the primers will perfectly match the unknown template. To digest the dsDNA into multiple small fragments, we used a cocktail of five 4-cutter enzymes that have different recognition sites but all leave GC-5′ overhangs. After the digestion, the DNA fragments were treated with Klenow DNA polymerase in the presence of dCTP to fill in the 5′ overhangs with 1 nt. This step converts the self-complementary 5′-GC protrusions into 5′-C overhangs that could ligate only to the synthetic adapters designed accordingly. To convert these short DNA fragments into primers that could be recovered and used in the enrichment protocol, we 'tagged' them by ligation to adapters (Figure 1, B2–B3). In the context of the human genome (3.2 Gb) (26), 18 nt is the minimum required length (x) for the creation of a specific oligonucleotide, calculated by the formula $Nx/4^x < 1$, where N is the size of the target. For a large viral genome (e.g. $N = 100$ Kb) this minimum length is reduced to 10 bp. However, if the aim is to distinguish a viral genome of that size with the the human genome as a background, we still need a minimum size of 18 bp to ensure unique sequence specificity. The PEER protocol uses adapters with MmeI site, thus generating primers with 20 bp of unique sequence derived from the target (Figure 1, B3). This feature and the fact that the protocol is theoretically able to generate up to 38 primer tags/1000 bp sequence provides ample specificity. The adapter can alternatively incorporate sites for other IIS restriction enzymes. We have already successfully used adapters with BpmI in our preliminary experiments. The IIS cleavage allows all cDNA fragments of various lengths that have acquired adapters to be 'trimmed' to a uniform length. This generates a population of molecules that are suitable for extension reactions at a reasonably narrow temperature range. Class IIS endonucleases that leave 5′-protruding ends (20) are not suitable for PEER because they will generate self-blocking primers. Additional steps to ensure elimination of the background amplification and increase specificity may include synthesis of the Tester 23 dsDNA in the presence of dUTP and subsequent uracil-N-glycosylase (UDG) treatment of the primer capture product (after step D1 in Figure 1).

As established, to find a high titer virus (e.g. 10^8) within a library representing the entire human genome one needs to only search through $\sim 10^2$ of the clones since $\sim$2–3% of this library should contain viral sequences, however if the viral titer is 10^3 one needs to screen 10^7 clones. To circumvent exhaustive screening of the low titer tester libraries by colony hybridization we evaluated the copy number of targets of interest in the dsDNA by PCR and spot hybridization instead. A PCR approach, although very reassuring when used on the dsDNA material prior to enrichment, cannot be applied on the PEER products because they may not be comprised of fragments that will contain both priming sites. To assess the presence of the targets of interest in the PEER products we did spot hybridizations and from the corresponding libraries isolated colonies at random and sequenced them. We believe that the observed higher enrichment values obtained for the lower titer library is attributed to the fact that dilution of the tester material with the driver pool creates close to perfect driver match and thus greater blocking efficiency.

PEER is a conceptually new approach for the subtractive enrichment of complex nucleic acid mixtures and represents a novel use for both class IIS restriction enzymes and di-deoxytermination. The use of the tester to generate both primers and template for the subsequent PCR steps contributes to the specificity and sensitivity of PEER in comparison with other subtraction approaches. Unlike other known subtraction techniques where the selection step involves the hybridization of long DNA molecules which are eventually used as PCR templates, the PEER blocking step (Figure 1c) uses short tester-specific DNA fragments and through hybridization and highly specific enzymatic extension selects the unique ones for PCR primers. Such hybridization kinetics are significantly less complex than those between long DNA molecules and this may also contribute to the efficiency of the PEER. In addition, because PEER was designed to create normalized re-amplifiable double-stranded starting material, the method is suitable for use with samples of limited volume and is very cost efficient, especially when compared with new high-throughput sequencing methods. We did not observe loss of integrity of the background DNA, i.e. no recombination or insertion/deletion events in the sequences that we have seen. We found some primer multimers among the clones, but we did not quantify them because the sequence data were filtered for background/vector/primer noise prior to analysis. In theory PEER can also allow for several rounds of enrichment, as do RDA and SSH, i.e. the final PEER product could be digested again with the GC cutter cocktail or any other restriction enzyme combination, adapter-tagged, cut with MmeI and blocked on the same driver or even on an alternative driver, depending on the experimental goals. We have not attempted such experiments. Although other uses of PEER are beyond the scope of this study, a review of the strategy identifies steps that can be modified to increase the versatility of the technique.

Our findings demonstrate that PEER is robust, can be applied to different targets and can detect nucleic acids of unknown sequence at very low concentrations. In our experiments, PEER was able to detect a fragment of interest at very low initial concentration, outperformed the commercially available SSH technique and in the case of the HBV test target it was successful without the availability of a perfect driver. Collectively, the data obtained in this study indicate that PEER is more efficient than any other reported subtractive method in recovering target nucleic acids from complex mixtures.

ACKNOWLEDGEMENTS

Funding to pay the Open Access publication charges for this article was provided by the Centers for Disease Control, USA.

Conflict of interest statement. None declared.

REFERENCES

1. Muerhoff,A.S., Leary,T.P., Desai,S.M. and Mushahwar,I.K. (1997) Amplification and subtraction methods and their application to the discovery of novel human viruses. *J. Med. Virol.*, **53**, 96–103.
2. Liang,P. and Pardee,A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, **257**, 967–971.
3. Matz,M.V. and Lukyanov,S.A. (1998) Different strategies of differential display: areas of application. *Nucleic Acids Res.*, **26**, 5537–5543.
4. Lisitsyn,N., Lisitsyn,N. and Wigler,M. (1993) Cloning the differences between two complex genomes. *Science*, **259**, 946–951.
5. Allander,T., Emerson,S.U., Engle,R.E., Purcell,R.H. and Bukh,J. (2001) A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc. Natl Acad. Sci. USA*, **98**, 11609–11614.
6. Birkenmeyer,L.G., Leary,T.P., Muerhoff,A.S., Dawson,G.J., Mushahwar,I.K. and Desai,S.M. (2003) Selectively primed adaptive driver RDA (SPAD-RDA): an improved method for subtractive hybridization. *J. Med. Virol.*, **71**, 150–159.
7. Luo,J.-H., Puc,J.A., Slosberg,E.D., Yao,Y., Bruce,J.N., WrightJr,T.C., Becich,M.J. and Parsons,R. (1999) Differential subtraction chain, a method for identifying differences in genomic DNA and mRNA. *Nucleic Acids Res.*, **27**, e24.
8. Lavery,D.J., Lopez-Molina,L., Fleury-Olela,F. and Schibler,U. (1997) Selective amplification via biotin- and restriction-mediated enrichment (SABRE), a novel selective amplification procedure for detection of differentially expressed mRNAs. *Proc. Natl Acad. Sci. USA*, **94**, 6831–6836.
9. Diatchenko,L., Lau,Y.F., Campbell,A.P., Chenchik,A., Moqadam,F., Huang,B., Lukyanov,S., Lukyanov,K., Gurskaya,N. and Sverdlov,E.D. (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl Acad. Sci. USA*, **93**, 6025–6030.
10. Drmanac,S. and Drmanac,R. (1994) Processing of cDNA and genomic kilobase-size clones for massive screening, mapping and sequencing by hybridization. *Biotechniques*, **17**, 328–336.
11. Schena,M., Heller,R.A., Theriault,T.P., Konrad,K., Lachenmeier,E. and Davis,R.W. (1999) Microarrays: biotechnology's discovery platform for functional genomics. *Trends Biotechnol.*, **16**, 217–218.
12. Wang,D., Urisman,A., Liu,Y.T., Springer,M., Ksiazek,T.G., Erdman,D.D., Mardis,E.R., Hickenbotham,M., Magrini,V. and Eldred,J. (2003) Viral discovery and sequence recovery using DNA microarrays. *PLoS. Biol.*, **1**, E2.
13. Agaton,C., Unneberg,P., Sievertzon,M., Holmberg,A., Ehn,M., Larsson,M., Odeberg,J., Uhlen,M. and Lundeberg,J. (2002) Gene expression analysis by signature pyrosequencing. *Gene*, **289**, 31–39.
14. Leamon,J.H., Lee,W.L., Tartaro,K.R., Lanza,J.R., Sarkis,G.J., deWinter,A.D., Berka,J. and Lohman,K.L. (2003) A massively parallel PicoTiterPlate based platform for discrete picoliter-scale polymerase chain reactions. *Electrophoresis*, **24**, 3769–3777.
15. Yang,G.P., Ross,D.T., Kuang,W.W., Brown,P.O. and Weigel,R.J. (1999) Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes. *Nucleic Acids Res.*, **27**, 1517–1523.
16. Ganova-Raeva,L., Smith,A.W., Fields,H. and Khudyakov,Y. (2004) New calicivirus isolated from walrus. *Virus Res.*, **102**, 207–213.
17. Birkenmeyer,L.G., Desai,S.M., Muerhoff,A.S., Leary,T.P., Simons,J.N., Montes,C.C. and Mushahwar,I.K. (1998) Isolation of a GB virus-related genome from a chimpanzee. *J. Med. Virol.*, **56**, 44–51.
18. Simons,J.N., Pilot-Matias,T.J., Leary,T.P., Dawson,G.J., Desai,S.M., Schlauder,G.G., Muerhoff,A.S., Erker,J.C., Buijk,S.L. and Chalmers,M.L. (1995) Identification of two flavivirus-like genomes in the GB hepatitis agent. *Proc. Natl Acad. Sci. USA*, **92**, 3401–3405.
19. Ji,W., Wright,M.B., Cai,L., Flament,A. and Lindpaintner,K. (2002) Efficacy of SSH PCR in isolating differentially expressed genes. *BMC Genomics*, **3**, 12.
20. Szybalsky,W., Kim,S.C., Hasan,N. and Podhajska,A.J. (1991) Class-IIS restriction enzymes—a review. *Gene*, **100**, 13–26.
21. Wu,J., Sullivan,D.E. and Gerber,M.A. (1994) Quantitative polymerase chain reaction for Hepatitis B virus DNA. *J. Virol. Methods*, **49**, 331–342.
22. Kotewicz,M.L., Sampson,C.M., D'Alessio,J.M. and Gerard,G.F. (1988) Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res.*, **16**, 265–277.
23. Boyd,A.C., Charles,I.G., Keyte,J.W. and Brammar,W.J. (1986) Isolation and computer-aided characterization of MmeI, a type II

restriction endonuclease from *Methylophilus methylotrophus*. *Nucleic Acids Res.*, **14**, 5255–5274.
24. Tucholski,J., Skowron,P.M. and Podhajska,A.J. (1995) MmeI, a class-IIS restriction endonuclease: purification and characterization. *Gene*, **157**, 87–92.
25. Seth,D., Gorrell,M.D., McGuinness,P.H., Leo,M.A., Lieber,C.S., McCaughan,G.W. and Haber,P.S. (2003) SMART amplification maintains representation of relative gene expression: quantitative validation by real time PCR and application to studies of alcoholic liver disease in primates. *J. Biochem. Biophys. Methods*, **55**, 53–66.
26. Wright,F.A., Lemon,W.J., Zhao,W.D., Sears,R., Zhuo,D., Wang,J.P., Yang,H.Y., Baer,T., Stredney,D. and Spitzner,J. (2001) A draft annotation and overview of the human genome. *Genome Biol.*, **2**, research0025.1–research0025.18.